

Bao LI

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/664,363

DATE: 03/27/2001
TIME: 23:23:35

INPUT SET: S36579.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

1
2
3 (1) General Information:
4
5 (i) APPLICANT: Highfield, Peter Edmund
6 Rodgers, Brian Colin
7 Tedder, Richard Seton
8 Barbara, John Anthony James
9
10 (ii) TITLE OF INVENTION: Viral Agent
11
12 (iii) NUMBER OF SEQUENCES: 25
13
14 (iv) CORRESPONDENCE ADDRESS:
15 (A) ADDRESSEE: Rothwell, Figg, Ernst & Kurz
16 (B) STREET: 1700 K Street
17 (C) CITY: Washington
18 (D) STATE: D.C.
19 (E) COUNTRY: U.S.A.
20 (F) ZIP: 20006
21
22 (v) COMPUTER READABLE FORM:
23 (A) MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage
24 (B) COMPUTER: IBM AT compatible
25 (C) OPERATING SYSTEM: MS-DOS V3.2
26 (D) SOFTWARE: Wordperfect 5.0 (DOS text)
27
28 (vi) CURRENT APPLICATION DATA:
29 (A) APPLICATION NUMBER: 09/664,363
30 (B) FILING DATE:
31 (C) CLASSIFICATION:
32
33 (vii) PRIOR APPLICATION DATA:
34 (A) APPLICATION NUMBER: 07/628,516
35 (B) FILING DATE: 17 DEC 1990
36
37 (A) APPLICATION NUMBER: UK 89 28 562.1
38 (B) FILING DATE: 18 DEC 1989
39
40 (vii) PRIOR APPLICATION DATA:
41 (A) APPLICATION NUMBER: UK 90 04 414.0
42 (B) FILING DATE: 27 FEB 1990
43
44 (vii) PRIOR APPLICATION DATA:
45 (A) APPLICATION NUMBER: UK 90 04 814.1
46 (B) FILING DATE: 03 MAR 1990

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/664,363DATE: 03/27/2001
TIME: 23:23:35

INPUT SET: S36579.raw

47
48 (viii) ATTORNEY/AGENT INFORMATION:
49 (A) NAME: E. Anthony Figg
50 (B) REGISTRATION NUMBER: 27,195
51 (C) REFERENCE/DOCKET NUMBER: 1645-103A
52
53 (xi) TELECOMMUNICATION INFORMATION:
54 (A) TELEPHONE: (202) 833-5740
55 (B) TELEFAX: (202) 833-5744
56
57
58 (2) INFORMATION FOR SEQ ID NO:1:
59
60 (i) SEQUENCE CHARACTERISTICS:
61 (A) LENGTH: 21 bases
62 (B) TYPE: nucleotide
63 (C) STRANDEDNESS: single
64 (D) TOPOLOGY: linear
65 (ii) MOLECULE TYPE: synthetic DNA
66
67 (vi) ORIGINAL SOURCE:
68 (A) ORGANISM: bacteriophage lambda gt11
69
70 (vii) IMMEDIATE SOURCE:
71 (A) LIBRARY: Oligonucleotide synthesizer; oligo d19
72
73 (ix) FEATURE:
74 (B) LOCATION: from 1 to 21 bases homologous to upstream portion
75 of lacZ gene flanking the EcoR1 site in bacteriophage
76 lambda gt11
77 (D) OTHER INFORMATION: primes DNA synthesis from the phage
78 vector into cDNA inserted at the EcoR1 site.
79
80 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
81
82 GGTGGCGACG ACTCCTGGAG C 21
83
84
85 (2) INFORMATION FOR SEQ ID NO:2:
86
87 (i) SEQUENCE CHARACTERISTICS:
88 (A) LENGTH: 21 bases
89 (B) TYPE: nucleotide
90 (C) STRANDEDNESS: single
91 (D) TOPOLOGY: linear
92
93 (ii) MOLECULE TYPE: synthetic DNA
94
95 (vi) ORIGINAL SOURCE:
96 (A) ORGANISM: bacteriophage lambda gt11
97
98 (vii) IMMEDIATE SOURCE:
99 (A) LIBRARY: Oligonucleotide synthesizer; oligo d20

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/664,363DATE: 03/27/2001
TIME: 23:23:35

INPUT SET: S36579.raw

100
101 (ix) FEATURE:
102 (B) LOCATION: from 1 to 21 bases homologous to downstream
103 portion of lacZ gene flanking the EcoR1 site in
104 bacteriophage lambda gt11
105 (D) OTHER INFORMATION: primes DNA synthesis from the phage
106 vector into cDNA inserted at the EcoR1
107 site.
108
109 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
110
111 TTGACACCAG ACCAACTGGT A 21
112
113
114 (2) INFORMATION FOR SEQ ID NO:3:
115
116 (i) SEQUENCE CHARACTERISTICS:
117 (A) LENGTH: 1770 base pairs
118 (B) TYPE: nucleotide with corresponding protein
119 (C) STRANDEDNESS: single
120 (D) TOPOLOGY: linear
121
122 (ii) MOLECULE TYPE: cDNA to genomic RNA
123
124 (vi) ORIGINAL SOURCE:
125 (A) ORGANISM: human; serum infectious for PT-NANBH
126
127 (vii) IMMEDIATE SOURCE:
128 (A) LIBRARY: clone JG2 from cDNA library in lambda gt11
129
130 (ix) FEATURE:
131 (B) LOCATION: from 1 to 1770 bp portion of the PT-NANBH
132 polyprotein
133 (D) OTHER INFORMATION: probably encodes viral non-structural
134 proteins
135
136 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
137
138 CAA AAT GAC TTC CCA GAC GCT GAC CTC ATC GAG GCC AAC CTC CTG TGG 48
139 Gln Asn Asp Phe Pro Asp Ala Asp Leu Ile Glu Ala Asn Leu Leu Trp
140 5 10 15
141
142 CGG CAT GAG ATG GGC GGG GAC ATT ACC CGC GTG GAG TCA GAG AAC AAG 96
143 Arg His Glu Met Gly Gly Asp Ile Thr Arg Val Glu Ser Glu Asn Lys
144 20 25 30
145
146 GTA GTA ATC CTG GAC TCT TTC GAC CCG CTC CGA GCG GAG GAG GAT GAG 144
147 Val Val Ile Leu Asp Ser Phe Asp Pro Leu Arg Ala Glu Glu Asp Glu
148 35 40 45
149
150 CGG GAA GTG TCC GTC CCG GCG GAG ATC CTG CGG AAA TCC AAG AAA TTC 192
151 Arg Glu Val Ser Val Pro Ala Glu Ile Leu Arg Lys Ser Lys Lys Phe
152 50 55 60

INPUT SET: S36579.raw

153																		
154	CCA	CCA	GCG	ATG	CCC	GCA	TGG	GCA	CGC	CCG	GAT	TAC	AAC	CCT	CCG	CTG	240	
155	Pro	Pro	Ala	Met	Pro	Ala	Trp	Ala	Arg	Pro	Asp	Tyr	Asn	Pro	Pro	Leu		
156	65					70					75					80		
157																		
158	CTG	GAG	TCC	TGG	AAG	GCC	CCG	GAC	TAC	GTC	CCT	CCA	GTG	GTA	CAT	GGG	288	
159	Leu	Glu	Ser	Trp	Lys	Ala	Pro	Asp	Tyr	Val	Pro	Pro	Val	Val	His	Gly		
160					85					90					95			
161																		
162	TGC	CCA	CTG	CCA	CCT	ACT	AAG	ACC	CCT	CCT	ATA	CCA	CCT	CCA	CGG	AGA	336	
163	Cys	Pro	Leu	Pro	Pro	Thr	Lys	Thr	Pro	Pro	Ile	Pro	Pro	Pro	Arg	Arg		
164				100					105					110				
165																		
166	AAG	AGG	ACA	GTT	GTT	CTG	ACA	GAA	TCC	ACC	GTG	TCT	TCT	GCC	CTG	GCG	384	
167	Lys	Arg	Thr	Val	Val	Leu	Thr	Glu	Ser	Thr	Val	Ser	Ser	Ala	Leu	Ala		
168			115					120					125					
169																		
170	GAG	CTT	GCC	ACA	AAG	GCT	TTT	GGT	AGC	TCC	GGA	CCG	TCG	GCC	GTC	GAC	432	
171	Glu	Leu	Ala	Thr	Lys	Ala	Phe	Gly	Ser	Ser	Gly	Pro	Ser	Ala	Val	Asp		
172		130					135					140						
173																		
174	AGC	GGC	ACG	GCA	ACC	GCC	CCT	CCT	GAC	CAA	TCC	TCC	GAC	GAC	GGC	GGA	480	
175	Ser	Gly	Thr	Ala	Thr	Ala	Pro	Pro	Asp	Gln	Ser	Ser	Asp	Asp	Gly	Gly		
176	145					150					155					160		
177																		
178	GCA	GGA	TCT	GAC	GTT	GAG	TCG	TAT	TCC	TCC	ATG	CCC	CCC	CTT	GAG	GGG	528	
179	Ala	Gly	Ser	Asp	Val	Glu	Ser	Tyr	Ser	Ser	Met	Pro	Pro	Leu	Glu	Gly		
180					165					170					175			
181																		
182	GAG	CCG	GGG	GAC	CCC	GAT	CTC	AGC	GAC	GGG	TCT	TGG	TCT	ACC	GTG	AGT	576	
183	Glu	Pro	Gly	Asp	Pro	Asp	Leu	Ser	Asp	Gly	Ser	Trp	Ser	Thr	Val	Ser		
184				180					185					190				
185																		
186	GAG	GAG	GCC	GGT	GAG	GAC	GTC	GTC	TGC	TGC	TCG	ATG	TCC	TAC	ACA	TGG	624	
187	Glu	Glu	Ala	Gly	Glu	Asp	Val	Val	Cys	Cys	Ser	Met	Ser	Tyr	Thr	Trp		
188			195					200					205					
189																		
190	ACA	GGC	GCT	CTG	ATC	ACG	CCA	TGC	GCT	GCG	GAG	GAA	AGC	AAG	CTG	CCC	672	
191	Thr	Gly	Ala	Leu	Ile	Thr	Pro	Cys	Ala	Ala	Glu	Glu	Ser	Lys	Leu	Pro		
192		210					215											

RAW SEQUENCE LISTING PATENT APPLICATION US/09/664,363

 DATE: 03/27/2001
 TIME: 23:23:36

INPUT SET: S36579.raw

206	ATG AAG GCG AAG GCG TCC ACA GTT AAG GCT AAG CTT CTA TCA GTA GAG	864
207	Met Lys Ala Lys Ala Ser Thr Val Lys Ala Lys Leu Leu Ser Val Glu	
208	275 280 285	
209		
210	GAA GCC TGC AAG CTG ACG CCC CCA CAT TCG GCC AAA TCT AAA TTT GGC	912
211	Glu Ala Cys Lys Leu Thr Pro Pro His Ser Ala Lys Ser Lys Phe Gly	
212	290 295 300	
213		
214	TAT GGG GCA AAG GAC GTC CGG AAC CTA TCC AGC AAG GCC ATT AAC CAC	960
215	Tyr Gly Ala Lys Asp Val Arg Asn Leu Ser Ser Lys Ala Ile Asn His	
216	305 310 315 320	
217		
218	ATC CGC TCC GTG TGG GAG GAC TTG TTG GAA GAC ACT GAA ACA CCA ATT	1008
219	Ile Arg Ser Val Trp Glu Asp Leu Leu Glu Asp Thr Glu Thr Pro Ile	
220	325 330 335	
221		
222	GAC ACC ACC ATC ATG GCA AAA AAT GAG GTT TTC TGC GTC CAA CCA GAG	1056
223	Asp Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Gln Pro Glu	
224	340 345 350	
225		
226	AGA GGA GGC CGC AAG CCA GCT CGC CTT ATC GTG TTC CCA GAC TTG GGG	1104
227	Arg Gly Gly Arg Lys Pro Ala Arg Leu Ile Val Phe Pro Asp Leu Gly	
228	355 360 365	
229		
230	GTC CGT GTG TGC GAG AAA ATG GCC CTC TAT GAC GTG GTC TCC ACC CTC	1152
231	Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Val Ser Thr Leu	
232	370 375 380	
233		
234	CCT CAG GCT GTG ATG GGC TCC TCG TAC GGA TTC CAG TAT TCT CCT GGA	1200
235	Pro Gln Ala Val Met Gly Ser Ser Tyr Gly Phe Gln Tyr Ser Pro Gly	
236	385 390 395 400	
237		
238	CAG CGG GTC GAG TTC CTG GTG AAC GCC TGG AAA TCA AAG AAG ACC CCT	1248
239	Gln Arg Val Glu Phe Leu Val Asn Ala Trp Lys Ser Lys Lys Thr Pro	
240	405 410 415	
241		
242	ATG GGC TTT GCA TAT GAC ACC CGC TGT TTT GAC TCA ACA GTC ACT GAG	1296
243	Met Gly Phe Ala Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu	
244	420 425 430	
245		
246	AAT GAC ATC CGT GTA GAG GAG TCA ATT TAT CAA TGT TGT GAC TTG GCC	1344
247	Asn Asp Ile Arg Val Glu Glu Ser Ile Tyr Gln Cys Cys Asp Leu Ala	
248	435 440 445	
249		
250	CCC GAA GCC AGA CAG GCC ATA AGG TCG CTC ACA GAG CGG CTT TAT ATC	1392
251	Pro Glu Ala Arg Gln Ala Ile Arg Ser Leu Thr Glu Arg Leu Tyr Ile	
252	450 455 460	
253		
254	GGG GGT CCC CTG ACT AAT TCA AAA GGG CAG AAC TGC GGC TAT CGC CGG	1440
255	Gly Gly Pro Leu Thr Asn Ser Lys Gly Gln Asn Cys Gly Tyr Arg Arg	
256	465 470 475 480	
257		
258	TGC CGC GCG AGC GGC GTG CTG ACG ACT AGC TGC GGT AAT ACC CTC ACA	1488

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/664,363

DATE: 03/27/2001
TIME: 23:23:36

INPUT SET: S36579.raw

Line

Error

Original Text

PAGE: 1

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION US/09/664,363

DATE: 03/27/2001
TIME: 23:23:36

INPUT SET: S36579.raw

< < THERE ARE NO ITEMS MISSING > >

PAGE: 1

SEQUENCE CORRECTION REPORT
PATENT APPLICATION US/09/664,363

DATE: 03/27/2001
TIME: 23:23:36

INPUT SET: S36579.raw

Line

Original Text

Corrected Text